Surviving in the wilderness.

James B. Wilson
Colorado State University

Wildness predicts the inevitable failure to catalog your objects.

Why did you want a catalog in the first place?

Researcher: My proof works except for groups of order 1536 and 2050. Those I'll do by hand.

```
| ...Magma available through Simons Foundation ...|

>NumberOfSmallGroups(1536);
408641062
> NumberOfSmallGroups(2050);
Runtime error:  The groups of order 2050 are not small
```

...good news, my new theorem just became a conjecture!

Moral: Researcher decisions are influenced by knowing the number of cases. Even rough estimates are helpful.

Researcher: I don't have time to referee this paper; I'm sure it is the same semifield I found last year anyway.

```
 >IsIsomorphic(A1, A2);
false
```

...OK, I'll just be picky about grammar instead.

Moral: Having practical tools to compare examples keeps you honest. – E. A. O'Brien

Researcher: There must be more examples of p-groups with G_2 as acentral automorphisms. What are they like?

```
>p := 7;
>G := my_GlasbyPalfySchneider_group(p);
>H := RandomSibling( G, change:=[ "Size" ],
 preserve:=["Nilpotence", "pClass", "Out"] );
```

Science fiction? No! But it will be difficult to implement.

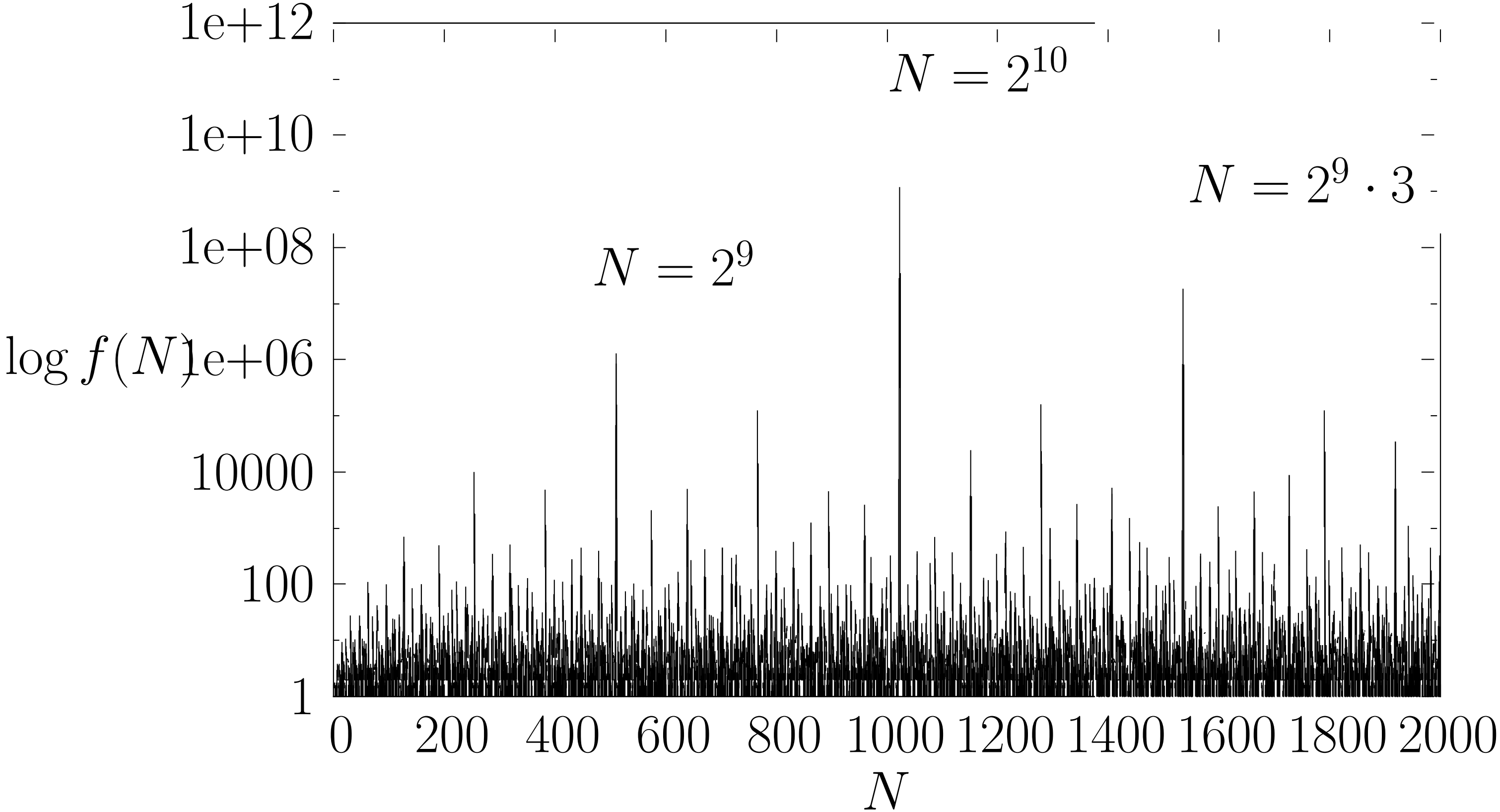
Survival after catalogs depends at least on these:

Counting: theorems, and eventually algorithms, to estimate quantities of objects that would have been in the catalog.
Comparing: tools to test appropriate equality.
Creating: methods to sample pseudo-randomly.

(Is this research or engineering? Maybe both, but who other than researchers could actually solve these?)

COUNTING IN ALGEBRA

Besche-Eick-O'Brien 2000.



A log-scale plot of the number $f(N)$ of the groups of order N.

(Probably) most finite groups order 2^k, 2^k3, 3^k....

Conjecture. Erdős
Up to isomorphism most groups of size $\leq N$ have order 2^k.

Theorem. Higman 60; Sims 65
The number $f(p^n)$ of groups of order p^n is

$$p^{2n^3/27+\Omega(n^2)\cap O(n^{3-\epsilon})}$$

for a some $\epsilon > 0$.

Theorem. Pyber 93 The number $f(N)$ of groups order at N satisfies

$$f(N) \leq N^{2\mu(N)^2/27+D\mu(N)^{2-\epsilon}}.$$

Fact. The number of graphs on N vertices is

$$2^{\Theta(N^2)}.$$

Fact. The number of semi-groups of order N vertices is

$$2^{\Theta(N^2 \log N)}.$$

Groups do not grow like combinatorics. The rare prime power sized sets are by far the most complex.

What grows like groups?

Theorem. Kruse-Price-70
The number of finite rings of order p^n is

$$p^{4n^3/27+\Omega(n^2)\cap O(n^{3-\epsilon})}$$

.

Theorem. Neretin-87
The dimension of the variety of algebras is

$$\frac{2}{27}n^3 + D_1 n^{3-\epsilon_1}$$

for commutative or Lie,

$$\frac{4}{27}n^3 + D_2 n^{3-\epsilon_2}$$

for associative.

Theorem. Poonen-08
The number of commutative rings of order p^n is

$$p^{2n^3/27+\Omega(n^2)\cap O(n^{3-\epsilon})}$$

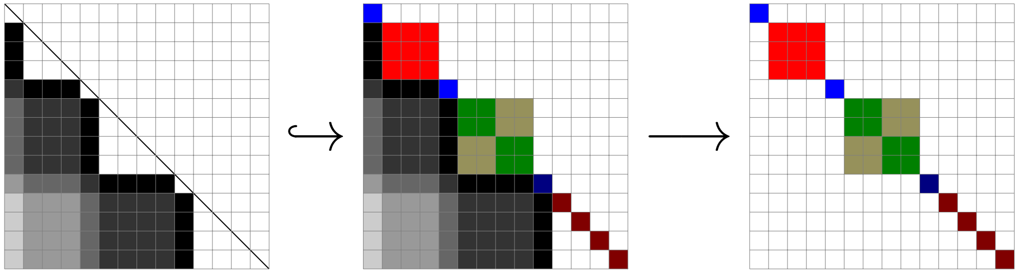
Why so similar to groups?
Hint.
Groups have a second product

$$[x, y] = x^{-1}x^y = x^{-1}y^{-1}xy$$

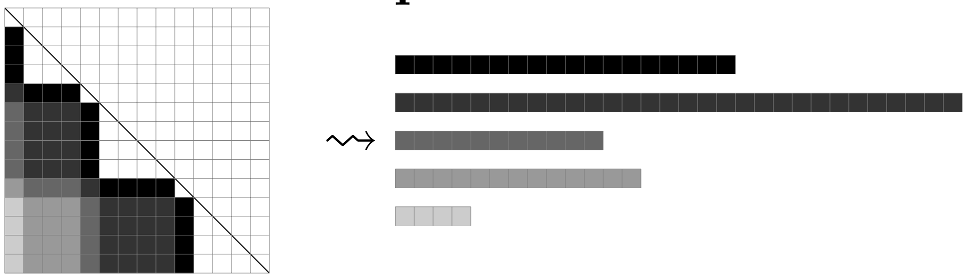
and it nearly distributes:

$$[xy, z] = [x, z]^y[y, z].$$

Step one: separate nilpotent from reductive

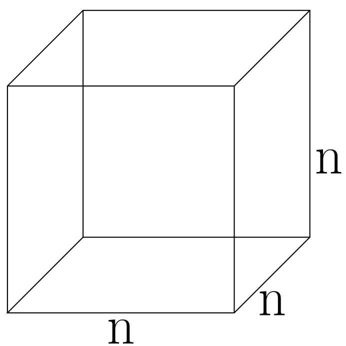


Step two: Break nilpotent into abelian sections



Where is the complexity in "triangular matrices"?

A. Nonassociative products need 3-dimensional array of parameters. Entropy of $\Theta(n^3)$.

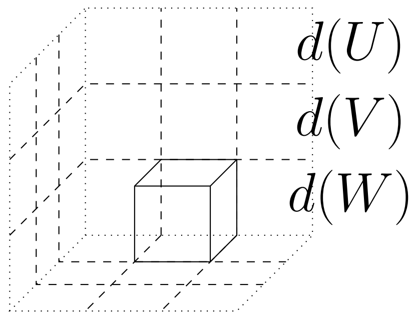


B. Matrix type groups

$$\begin{bmatrix} s & u & w \\ 0 & s & v \\ 0 & 0 & s \end{bmatrix} \begin{bmatrix} s' & u' & w' \\ 0 & s' & v' \\ 0 & 0 & s' \end{bmatrix} =$$

$$\begin{bmatrix} ss' & us'+su' & ws'+u*v'+sw' \\ 0 & ss' & vs'+sv' \\ 0 & 0 & ss' \end{bmatrix}$$

need only $* : U \times V \rightarrowtail W$.

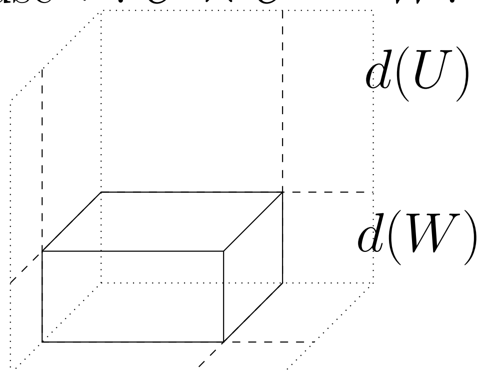


$$d(U)d(V)d(W) \leq n^3/27$$

C. Cut to diagonal embedding

$$\left\{ \begin{bmatrix} s & u & w \\ 0 & s & \pm u\theta \\ 0 & 0 & s \end{bmatrix} : \begin{matrix} u \in U, \\ w \in W \end{matrix} \right\}$$

now use $* : U \times U \rightarrowtail W$.



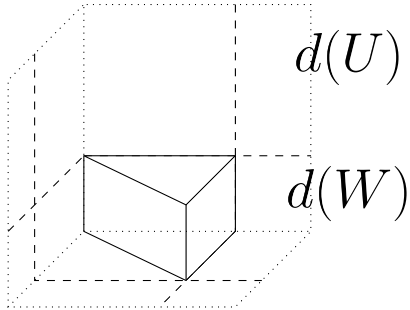
$$d(U)^2(n - d(U)) \le 4n^3/27.$$

D. Add symmetry

$$\left\{ \begin{bmatrix} s & u & w \\ 0 & s & \pm u\theta \\ 0 & 0 & s \end{bmatrix} : u \in U, w \in W \right\}$$

need $\pm\theta$-Hermitian

$$* : U \times U \rightarrowtail W.$$



$$\tfrac{1}{2} d(U)^2(n - d(U)) \le 2n^3/27.$$

CREATING IN ALGEBRA

Obvious default random sample.

To create a "random" group, ring, or algebra, we can just fill in the data structures we described in our counting.

Issues.

(1) Not all substitutions are consistent with group laws.
(2) Tends to give p-groups with probability 1. While "true", users want something different at times.

Other models...

Def.(Gromov '87-'03) $G = < x_1, \ldots, x_n | r_1, \ldots, r_s \rangle$ $|r_i| \geq \ell$ uniformly random (later models replace this 0-density with $\delta \in [0,1]$ density).

Theorem (Gromov) These groups are 1, $\mathbb{Z}/2$, or infinite hyperbolic (Cayley graph is tree like).

Theorem (Champetier '00) No measureable

$$f : \{\langle X | R \rangle\}/\cong \to \mathbb{R}.$$

(Doesn't play nice with isomorphism classes.)

Pick a random subgroup of a finite group?

Theorem (Dixon; Kantor-Lubotzky; Liebeck-Shalev) For A_n, S_n, and all groups of Lie type, two random elements generate with high probability.

(Mann) Try parabolic?

Theorem (W.) $U_d(q)$ upper uni-triangular matrices.
If $e > 2\sqrt{d}$ then sampling in U_d, then almost always

$$q^{d-e}|\langle u_1, \ldots, u_e\rangle| = |U_d|$$

in fact $U_d' = \langle u_1, \ldots, u_e\rangle'$.

(Probably) similar claims for all groups of Lie type and S_n.

No known "big groups" let you sample interesting random subgroups by generators.

Proof. Sims rank of a bimaps $* : A\times A \rightarrowtail B$ smallest dimension subspace $X \leq A$ where $A * A = B$.

In U_d commutation has Sims rank $\lceil\sqrt{d}\rceil$.

Prove generic $2\sqrt{d}$ subspace X of A satisfies $X * X = B$.
$\square$

This problem prevents useful random sampling of rings and algebras also.

A working but confusing heuristic.

Randomly sample **sparse** matrices $x_1, \ldots, x_e$ and the

$$\log_p |\langle x_1, \ldots, x_e \rangle|$$

becomes normally distributed.

So out of *less randomness* you get *more randomness*.

Does this make sense in theory?

Seems to be because this way $[x_i, x_j]$ are trivial often; so, generic large subspaces of sparse matrices avoid Sims subspaces.

Prove this or explain some other way.

HARD COUNTING AND CREATING IN ALGEBRA

Goal: Random sample from within a class $\mathcal{L}$ but also satisfy a property Q (equiv. $\neg Q$).

Assume $\mathcal{L} \subset \Sigma^* = \bigcup_i \Sigma^i$ a set of strings over an alphabet Σ.

Set $\mathcal{L}_Q = \{w \in \mathcal{L} : Q(w)\}$.

Def. For a language $\mathcal{L}$, a *padding* (p, q) are poly-time computable $p : \Sigma^* \times \Sigma^* \to \Sigma^*$, $q : \Sigma^* \to \Sigma^*$ such that

$$p(w, u) \in \mathcal{L} \Leftrightarrow w \in \mathcal{L},$$

$$q(p(w, u)) = u.$$

Pump randomness into language using <u>one</u> instance!

Prop. For a paddable $\mathcal{L}$, $\exists a, c > 0$,

$$\delta(n) = \frac{\mathcal{L} \cap \Sigma^n}{2^n} \in \Omega\left(a^{n^{1/c}}\right).$$

$c = 1$ if $|f(x, y)| \in O(|x|+|y|)$.

Coro.[Miyazaki-W.] $\mathcal{L}_Q$ has exponential density if it has linear padding. Also, gives a dense polynomial-time random sampling method.

Who has a padding?

All known NP-complete problems have linear paddings.

Conj. (Berman-Hartmanis) All NP-complete problems are isomorphic (bijective reductions).

Thm (Berman-Hartmanis) Paddable language that are poly-equivalent are isomorphic.

Thm (Miyazaki-W.) Linear paddable language that are linear-equivalent are linearly isomorphic.

Problem? `SAT` $\leq$ `CLIQUE` non-linear.

Fix. If *efficiently encoded* `SAT` $\leq$ `CLIQUE` can be made linear. (True of all NP-complete reductions tried.)

So we indeed expect NP-complete problems to have linear paddings.

Thm[M-W]. $\texttt{DV} - \texttt{SAT} \leq_{lin} \mathcal{L}$, and $\mathcal{L} \in NP$ having verifier V

- V 2-tape with RAM
- Oblivious computation.

Then $\mathcal{L}$ is *linear isomorphic* to $\texttt{DV-SAT}$, and *linearly paddable*. Hence $\texttt{DV-SAT}$ is linearly complete amongst these NP problems.

Ex. $\texttt{DV-SAT}$, $\texttt{AFF\_PT}_k$, $\texttt{SDIT}_k$, $\texttt{MINRANK}_k$, $\texttt{SINGULAR}_k$

Coro(Hard Counting). Problems complete for this class are dense.

Coro(Hard Sampling) Problems complete for this class have a polynomial time random sample algorithm needing one seed.

Def $* : U \times V \rightarrowtail W$ is *non-singular* if $u * v = 0$ implies $u = 0$ or $v = 0$. The *(left) singularity radical* is $R = \langle v : \exists u \neq 0, u * v = 0 \rangle$.

Thm[M-W]. Singular products a dense amongst general products.

1st proved by making linear reduction to NP-complete, then "demystified" to concrete proof.

Prob. Fix $M_i \in M_d(k)$. Decide if

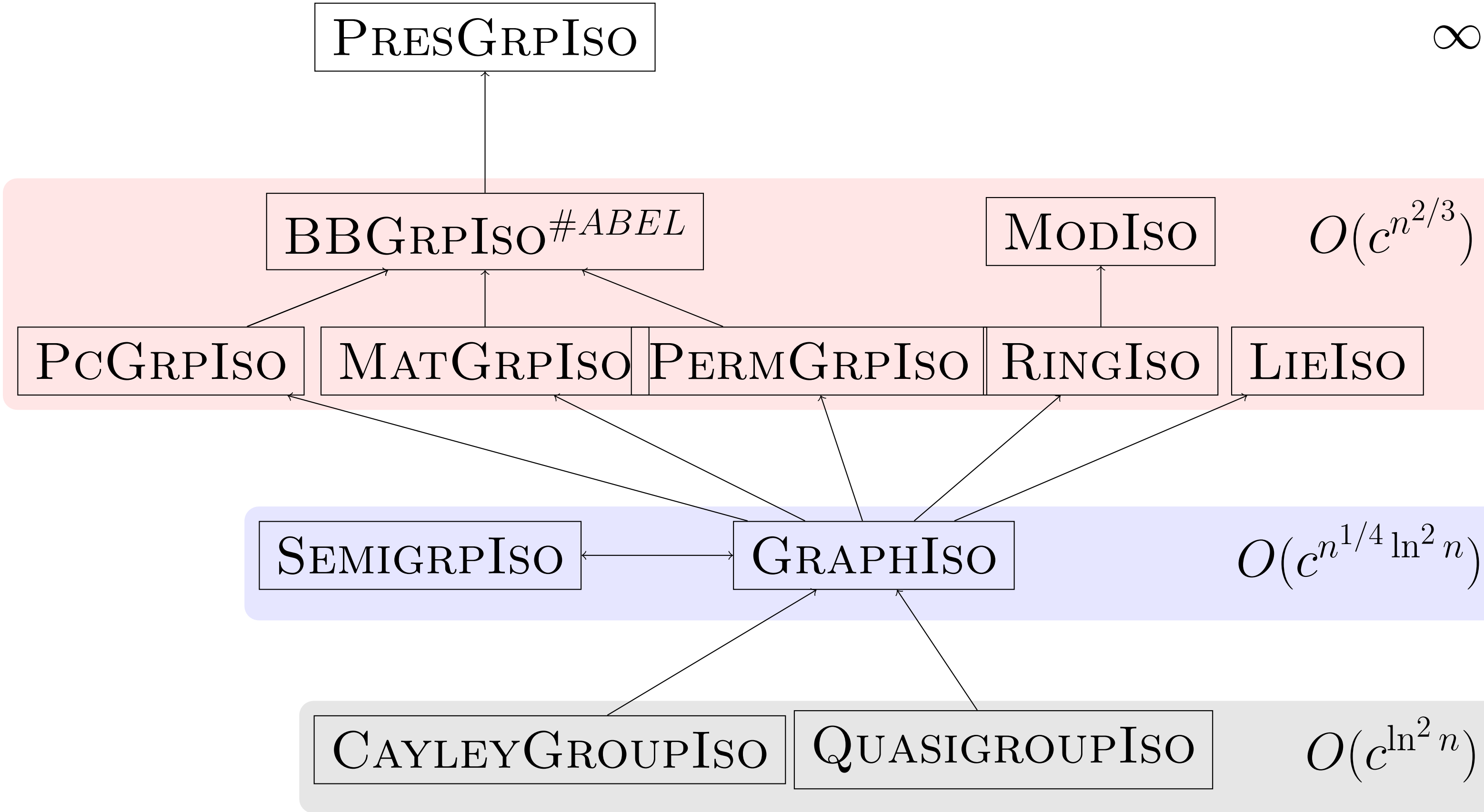
$$\det(x_1 M_1 + \cdots + x_d M_d) = 0$$

has a (projective) point.

If this problem is in our class then, there are exponentially many finite projective planes (an open conjecture studied by Albert, Knuth, Kantor, and many others).

COMPARING IN ALGEBRA

Isomorphism problems in algebra today.



$n =$ input size, e.g. graphs on v vertices have $n \in O(v^2)$

This is a long story, check out:
www.math.colostate.edu/∼jwilson/papers/group-iso-2015.pdf